

19  
Appendix 1

GCTATTGTCG ACGTATGACG TTTGCTCTAT AGCCATCGCT GCTCCCATGC GCGCCACTCG 60  
 GTCGCAGGGG GTGTGGGATT TTTTTTGGGA GACAATCCTC ATGGCCTATA CGACGGCCCA 120  
 GTTGGTGACT GCGTACACCA ACGCCAACCT CGGCAAGGCG CCTGACGCCG CCACCACGCT 180  
 GACGCTCGAC GCGTACGCGA CTCAAACCCA GACGGGCGGC CTCTCGGACG CCGCTGCGCT 240  
 GACCAACACC CTGAAGCTGG TCAACAGCAC GACGGCTGTT GCCATCCAGA CCTACCAGTT 300  
 CTTCAACGGC GTTGCCCCGT CGGCCGCTGG TCTGGACTTC CTGGTCGACT CGACCACCAA 360  
 CACCAACGAC CTGAACGACG CGTACTACTC GAAGTTCGCT CAGGAAAACC GCTTCATCAA 420  
 CTTCTCGATC AACCTGGCCA CGGGCGCCGG CGCCGGCGCG ACGGCTTTCG CCGCCGCCTA 480  
 CACGGGCGTT TCGTACGCC AGACGGTCGC CACCGCCTAT GACAAGATCA TCGGCAACGC 540  
 CGTCGCGACC GCCGCTGGCG TCGACGTCGC GGCCGCCGTG GCTTTCCTGA GCCGCCAGGC 600  
 CAACATCGAC TACCTGACCG CCTTCGTGCG CGCCAACACG CCGTTCACGG CCGCTGCCGA 660  
 CATCGATCTG GCCGTCAAGG CCGCCCTGAT CGGCACCATC CTGAACGCCG CCACGGTGTC 720  
 GGGCATCGGT GGTACGCGA CCGCCACGGC CGCGATGATC AACGACCTGT CGGACGGCGC 780  
 CCTGTGACG GACAACGCG CTGGCGTGAA CCTGTTCACC GCCTATCCGT CGTCGGGCGT 840  
 GTCGGGTTCG ACCCTCTCGC TGACCACCGG CACCGACACC CTGACGGGCA CCGCCAACAA 900  
 CGACACGTTT GTTGGGGTG AAGTCGCCGG CGCTGCGACC CTGACCGTTG GCGACACCCT 960  
 GAGCGGCGGT GCTGGCACCG ACGTCCTGAA CTGGGTGCAA GCTGCTGCGG TTACGGCTCT 1020  
 GCCGACCGGC GTGACGATCT CGGGCATCGA AACGATGAAC GTGACGTCGG GCGCTGCGAT 1080  
 CACCCTGAAC ACGTCTTCGG GCGTGACGGG TCTGACCGCC CTGAACACCA ACACCAGCGG 1140  
 CGCGGCTCAA ACCGTCACCG CCGGCGCTGG CCAGAACCTG ACCGCCACGA CCGCCGCTCA 1200  
 AGCCGCGAAC AACGTCGCCG TCGACGGGCG CGCCAACGTC ACCGTCGCCT CGACGGGCGT 1260  
 GACCTCGGGC ACGACCACGG TCGGCGCCAA CTCGGCCGCT TCGGGACCG TGTCGGTGAG 1320  
 CGTCGCGAAC TCGAGCACGA CCACCACGGG CGCTATCGCC GTGACCGGTG GTACGGCCGT 1380  
 GACCGTGGCT CAAACGGCCG GCAACGCCGT GAACACCACG TTGACGCAAG CCGACGTGAC 1440  
 CGTGACCGGT AACTCCAGCA CCACGGCCGT GACGGTCACC CAAACGGCCG CCGCCACCGC 1500  
 CGGCGCTACG GTCGCCGGTC GCGTCAACGG CGCTGTGACG ATCACCGACT CTGCCGCCG 1560  
 CTCGGCCACG ACCGCCGGCA AGATCGCCAC GGTCACCCTG GGCAGCTTCG GCGCCGCCAC 1620  
 GATCGACTCG AGCGCTCTGA CGACCGTCAA CCTGTGCGGC ACGGGCACCT CGCTCGGCAT 1680

007474 04260

20  
Appendix 1 (cont'd)

CGGCGCGGGC	GCTCTGACCG	CCACGCGGAC	CGCCAACACC	CTGACCCTGA	ACGTCAATGG	1740
TCTGACGACG	ACCGGCGCGA	TCACGGACTC	GGAAGCGGGT	GCTGACGATG	GTTTCACCAC	1800
CATCAACATC	GCTGGTTTGA	CCGCCTCTTC	GACGATCGCC	AGCCTGGTGG	CCGCCGACGC	1860
GACGACCCTG	AACATCTCGG	GCGACGCTCG	CGTCACGATC	ACCTCGCACA	CCGCTGCCGC	1920
CCTGACGGGC	ATCACGGTGA	CCAACAGCGT	TGGTGCGACC	CTCGGCGCCG	AACTGGCGAC	1980
CGGTCTGGTC	TTACGGGGCG	GCGCTGGCCG	TGACTCGATC	CTGCTGGGCG	CCACGACCAA	2040
GGCGATCGTC	ATGGGCGCCG	GCGACGACAC	CGTCACCGTC	AGCTCGGCGA	CCCTGGGCGC	2100
TGGTGGTTTC	GTCAACGGCG	GCGACGGCAC	CGACGTTCTG	GTGGCCAACG	TCAACGGTTC	2160
GTCGTTTACG	GCTGACCCGG	CCTTCGGCGG	CTTCGAAACC	CTCCGCGTCG	CTGGCGCGGC	2220
GGCTCAAGGC	TGGCACAACG	CCAACGGCTT	CACGGCTCTG	CAACTGGGCG	CGACGGCGGG	2280
TGGCAGGACC	TTCAACCAACG	TTGCGGTGAA	TGTCGGCCTG	ACCGTTCTGG	CGGCTCCGAC	2340
CGGTACGACG	ACCGTGACCC	TGGCCAACGC	CACGGGCACC	TCGGACGTGT	TCAACCTGAC	2400
CCTGTCTGTC	TGGGCGGCTC	TGGCCGCTGG	TACGGTTGCG	CTGGCTGGCG	TCGAGACGGT	2460
GAACATCGCC	GCCACCGACA	CCAACACGAC	CGCTCACGTC	GACACGCTGA	CGCTGCAAGC	2520
CACCTCGGCC	AAGTCGATCG	TGGTGACGGG	CAACGCCGGT	CTGAACCTGA	CCAACACCGG	2580
CAACACGGCT	GTCACCAAGT	TCGACGCCAG	CGCCGTACCC	GGCAGGGCTC	CGGCTGTGAC	2640
CTTCGTGTCT	GCCAACACCA	CGGTGGGTGA	AGTCGTCACG	ATCCGCGGCG	GCGCTGGCGC	2700
CGACTCGCTG	ACCGGTTTCG	CCACCGCCAA	TGACACCATC	ATCGGTGGCG	CTGGCGCTGA	2760
CACCTTGGTC	TACACCGGCG	GTACGGACAC	CTTCACGGGT	GGCACGGGCG	CGGATATCTT	2820
CGATATCAAC	GCTATCGGCA	CCTCGACCGC	TTTCGTGACG	ATCACCAGCG	CCGCTGTCTG	2880
CGACAAGCTC	GACCTCGTCG	GCATCTCGAC	GAACGGCGCT	ATCGCTGACG	GCGCCTTCGG	2940
CGCTGCGGTC	ACCCTGGGCG	CTGCTGCGAC	CCTGGCTCAG	TACCTGGACG	CTGCTGCTGC	3000
CGGCGACGGC	AGCGGCACCT	CGGTTGCCAA	GTGGTTCCAG	TTGGGCGGCG	ACACCTATGT	3060
CGTCGTTGAC	AGCTCGGCTG	GCGCGACCTT	CGTCAGCGGC	GCTGACGCGG	TGATCAAGCT	3120
GACCGGTCTG	GTCACGCTGA	CCACCTCGGC	CTTCGCCACC	GAAGTCCTGA	CGCTCGCCTA	3180
AGCGAACGTC	TGATCCTCGC	CTAGGCGAGG	ATCGCTAGAC	TAAGAGACCC	CGTCTTCCGA	3240
AAGGGAGGCG	GGGTCTTTCT	TATGGGCGCT	ACGCGCTGGC	CGGCCTTGCC	TAGTTCCGGT	3300

T05240" T24460

21  
Appendix 1 (cont'd)

Met Ala Tyr Thr Thr Ala Gln Leu Val Thr Ala Tyr Thr Asn Ala Asn  
1 5 10 15

Leu Gly Lys Ala Pro Asp Ala Ala Thr Thr Leu Thr Leu Asp Ala Tyr  
20 25 30

Ala Thr Gln Thr Gln Thr Gly Gly Leu Ser Asp Ala Ala Ala Leu Thr  
35 40 45

Asn Thr Leu Lys Leu Val Asn Ser Thr Thr Ala Val Ala Ile Gln Thr  
50 55 60

Tyr Gln Phe Phe Thr Gly Val Ala Pro Ser Ala Ala Gly Leu Asp Phe  
65 70 75 80

Leu Val Asp Ser Thr Thr Asn Thr Asn Asp Leu Asn Asp Ala Tyr Tyr  
85 90 95

Ser Lys Phe Ala Gln Glu Asn Arg Phe Ile Asn Phe Ser Ile Asn Leu  
100 105 110

Ala Thr Gly Ala Gly Ala Gly Ala Thr Ala Phe Ala Ala Ala Tyr Thr  
115 120 125

Gly Val Ser Tyr Ala Gln Thr Val Ala Thr Ala Tyr Asp Lys Ile Ile  
130 135 140

Gly Asn Ala Val Ala Thr Ala Ala Gly Val Asp Val Ala Ala Ala Val  
145 150 155 160

Ala Phe Leu Ser Arg Gln Ala Asn Ile Asp Tyr Leu Thr Ala Phe Val  
165 170 175

Arg Ala Asn Thr Pro Phe Thr Ala Ala Ala Asp Ile Asp Leu Ala Val  
180 185 190

Lys Ala Ala Leu Ile Gly Thr Ile Leu Asn Ala Ala Thr Val Ser Gly  
195 200 205

Ile Gly Gly Tyr Ala Thr Ala Thr Ala Ala Met Ile Asn Asp Leu Ser  
210 215 220

Asp Gly Ala Leu Ser Thr Asp Asn Ala Ala Gly Val Asn Leu Phe Thr  
225 230 235 240

Ala Tyr Pro Ser Ser Gly Val Ser Gly Ser Thr Leu Ser Leu Thr Thr  
245 250 255

Gly Thr Asp Thr Leu Thr Gly Thr Ala Asn Asn Asp Thr Phe Val Ala  
260 265 270

Gly Glu Val Ala Gly Ala Ala Thr Leu Thr Val Gly Asp Thr Leu Ser  
275 280 285

Gly Gly Ala Gly Thr Asp Val Leu Asn Trp Val Gln Ala Ala Ala Val  
290 295 300

Thr Ala Leu Pro Thr Gly Val Thr Ile Ser Gly Ile Glu Thr Met Asn  
305 310 315 320

Val Thr Ser Gly Ala Ala Ile Thr Leu Asn Thr Ser Ser Gly Val Thr  
325 330 335

Gly Leu Thr Ala Leu Asn Thr Asn Thr Ser Gly Ala Ala Gln Thr Val  
340 345 350

Sequence of the protein

22  
Appendix 1 (cont'd)

Thr Ala Gly Ala Gly Gln Asn Leu Thr Ala Thr Thr Ala Ala Gln Ala  
355 360 365

Ala Asn Asn Val Ala Val Asp Gly Arg Ala Asn Val Thr Val Ala Ser  
370 375 380

Thr Gly Val Thr Ser Gly Thr Thr Thr Val Gly Ala Asn Ser Ala Ala  
385 390 395 400

Ser Gly Thr Val Ser Val Ser Val Ala Asn Ser Ser Thr Thr Thr Thr  
405 410 415

Gly Ala Ile Ala Val Thr Gly Gly Thr Ala Val Thr Val Ala Gln Thr  
420 425 430

Ala Gly Asn Ala Val Asn Thr Thr Leu Thr Gln Ala Asp Val Thr Val  
435 440 445

Thr Gly Asn Ser Ser Thr Thr Ala Val Thr Val Thr Gln Thr Ala Ala  
450 455 460

Ala Thr Ala Gly Ala Thr Val Ala Gly Arg Val Asn Gly Ala Val Thr  
465 470 475 480

Ile Thr Asp Ser Ala Ala Ala Ser Ala Thr Thr Ala Gly Lys Ile Ala  
485 490 495

Thr Val Thr Leu Gly Ser Phe Gly Ala Ala Thr Ile Asp Ser Ser Ala  
500 505 510

Leu Thr Thr Val Asn Leu Ser Gly Thr Gly Thr Ser Leu Gly Ile Gly  
515 520 525

Arg Gly Ala Leu Thr Ala Thr Pro Thr Ala Asn Thr Leu Thr Leu Asn  
530 535 540

Val Asn Gly Leu Thr Thr Thr Gly Ala Ile Thr Asp Ser Glu Ala Ala  
545 550 555 560

Ala Asp Asp Gly Phe Thr Thr Ile Asn Ile Ala Gly Ser Thr Ala Ser  
565 570 575

Ser Thr Ile Ala Ser Leu Val Ala Ala Asp Ala Thr Thr Leu Asn Ile  
580 585 590

Ser Gly Asp Ala Arg Val Thr Ile Thr Ser His Thr Ala Ala Ala Leu  
595 600 605

Thr Gly Ile Thr Val Thr Asn Ser Val Gly Ala Thr Leu Gly Ala Glu  
610 615 620

Leu Ala Thr Gly Leu Val Phe Thr Gly Gly Ala Gly Arg Asp Ser Ile  
625 630 635 640

Leu Leu Gly Ala Thr Thr Lys Ala Ile Val Met Gly Ala Gly Asp Asp  
645 650 655

Thr Val Thr Val Ser Ser Ala Thr Leu Gly Ala Gly Gly Ser Val Asn  
660 665 670

Gly Gly Asp Gly Thr Asp Val Leu Val Ala Asn Val Asn Gly Ser Ser  
675 680 685

Phe Ser Ala Asp Pro Ala Phe Gly Gly Phe Glu Thr Leu Arg Val Ala  
690 695 700

FOUO" F22260

23  
Appendix 1. (cont'd)

Gly Ala Ala Ala Gln Gly Ser His Asn Ala Asn Gly Phe Thr Ala Leu  
705 710 715 720

Gln Leu Gly Ala Thr Ala Gly Ala Thr Thr Phe Thr Asn Val Ala Val  
725 730 735

Asn Val Gly Leu Thr Val Leu Ala Ala Pro Thr Gly Thr Thr Thr Val  
740 745 750

Thr Leu Ala Asn Ala Thr Gly Thr Ser Asp Val Phe Asn Leu Thr Leu  
755 760 765

Ser Ser Ser Ala Ala Leu Ala Ala Gly Thr Val Ala Leu Ala Gly Val  
770 775 780

Glu Thr Val Asn Ile Ala Ala Thr Asp Thr Asn Thr Thr Ala His Val  
785 790 795 800

Asp Thr Leu Thr Leu Gln Ala Thr Ser Ala Lys Ser Ile Val Val Thr  
805 810 815

Gly Asn Ala Gly Leu Asn Leu Thr Asn Thr Gly Asn Thr Ala Val Thr  
820 825 830

Ser Phe Asp Ala Ser Ala Val Thr Gly Thr Ala Pro Ala Val Thr Phe  
835 840 845

Val Ser Ala Asn Thr Thr Val Gly Glu Val Val Thr Ile Arg Gly Gly  
850 855 860

Ala Gly Ala Asp Ser Leu Thr Gly Ser Ala Thr Ala Asn Asp Thr Ile  
865 870 875 880

Ile Gly Gly Ala Gly Ala Asp Thr Leu Val Tyr Thr Gly Gly Thr Asp  
885 890 895

Thr Phe Thr Gly Gly Thr Gly Ala Asp Ile Phe Asp Ile Asn Ala Ile  
900 905 910

Gly Thr Ser Thr Ala Phe Val Thr Ile Thr Asp Ala Ala Val Gly Asp  
915 920 925

Lys Leu Asp Leu Val Gly Ile Ser Thr Asn Gly Ala Ile Ala Asp Gly  
930 935 940

Ala Phe Gly Ala Ala Val Thr Leu Gly Ala Ala Ala Thr Leu Ala Gln  
945 950 955 960

Tyr Leu Asp Ala Ala Ala Ala Gly Asp Gly Ser Gly Thr Ser Val Ala  
965 970 975

Lys Trp Phe Gln Phe Gly Gly Asp Thr Tyr Val Val Val Asp Ser Ser  
980 985 990

Ala Gly Ala Thr Phe Val Ser Gly Ala Asp Ala Val Ile Lys Leu Thr  
995 1000 1005

Gly Leu Val Thr Leu Thr Thr Ser Ala Phe Ala Thr Glu Val Leu Thr  
1010 1015 1020

Leu Ala  
1025

24  
Appendix 2

GAA TTC AGA TCT CAG GGC GCG GGG CAG GGT GGC TAT GGT GGG CTC GGC  
TCG CAA GGC  
GCT  
E F R S Q G A G Q G G Y G G L G S Q G A  
GGC CTG GGT GGC CAG GGC GCT GGC GCG GCC GCG GCC GCT GCG GCC GGT  
GGC  
G R G G Q G A G A A A A A A G G  
GCT GGC CAG GGC GGG CTG GGC TCG CAG GGC GCC GGC CAA GGC GCT GGC  
GCC GCG GCC  
GCT  
A G Q G G L G S Q G A G Q G A G A A A A  
GCG GCC GGT GGC GCC GGC CAG GGT GGC TAC GGC GGC CTG GGC AGC CAG  
GGC GCC GGT  
CGC  
A A G G A G Q G G Y G G L G S Q G A G R  
GGC GGT CAG GGC GCC GGT GCC GCG GCC GCT GCG GCC GGT GGC GCT GGG  
CAA GGC GGC TAC  
G G Q G A G A A A A A A G G A G Q G G Y  
GGC GGT CTG GGA TCC  
G G L G S

## Appendix 3

1/1

atg aac aca aac aag gca acc gca act tac ttg aaa tcc att atg ctt cca gag acc  
gga  
Met asn thr asn lys ala thr ala thr tyr leu lys ser ile met leu pro glu thr  
gly  
61/21

cca gca agc atc ccg gac gac ata acg gag aga cac atc tta aaa caa gag acc tgg  
tca  
pro ala ser ile pro asp asp ile thr glu arg his ile leu lys gln glu thr ser  
ser  
121/41

tac aac tta gag gtc tcc gaa tca gga agt ggc att ctt gtt tgt ttc cct ggg gca  
cca  
tyr asn leu glu val ser glu ser gly ser gly ile leu val cys phe pro gly ala  
pro  
181/61

ggc tca cgg atc ggt gca cac tac aga tgg aat grg aac cag acg ggg ctg gag ttc  
gac  
gly ser arg ile gly ala his tyr arg trp asn ala asn gln thr gly leu glu phe  
asp  
241/81

cag tgg ctg gag acg tgg cag gac ctg aag aaa gcc ttc aac tac ggg agg ctg atc  
tca  
gln trp leu glu thr ser gln asp leu lys lys ala phe asn tyr gly arg leu ile  
ser  
301/101

agg aaa tac gac att caa agc tcc aca cta ccg gcc ggt ctc tat gct ctg aac ggg  
acg  
arg lys tyr asp ile gln ser ser thr leu pro ala gly leu tyr ala leu asn gly  
thr  
361/121

ctc aac gct gcc acc ttc gaa ggc agt ctg tct gag gtg gag agc ctg acc tac aat  
agc  
leu asn ala ala thr phe glu gly ser leu ser glu val glu ser leu thr tyr asn  
ser  
421/141

ctg atg tcc cta act acg aac ccc cag gac aaa gcc aac aac cag ctg gtc acc aaa  
gga  
leu met ser leu thr thr asn pro gln asp lys ala asn asn gln leu val thr lys  
gly  
481/161

gtc acc gtc ctg aat cta cca aca ggg ttc gac aaa cca tac gtc cgc cta gag gac  
gag  
val thr val leu asn leu pro thr gly phe asp lys pro tyr val arg leu glu asp  
glu  
541/181

aca ccc cag ggt ctc cag tca atg aac ggg gcc agg atg agg tgc aca gct cca att  
gca  
thr pro gln gly leu gln ser met asn gly ala arg met arg cys thr ala ala ie  
ala  
601/201

cca cgg agg tac gag atc gac ctc cca tcc caa agc cta ccc ccc gtt cct ggg aca  
gga  
pro arg arg tyr glu ile asp leu pro ser gln ser leu pro pro val pro ala trp  
gly  
661/221

acc ctc acc act ctc tac gag gga aac gcc gac atc gtc agc tcc aca aca gtc acg  
gga  
thr leu thr thr leu tyr glu gly asn ala asp ile val ser ser thr thr val thr  
gly  
721/241

gac ala aac ttc agt ctg gca gaa cga ccc gca aac gag acc agg ttc gac tc cag  
ctg  
asp ile asn phe ser leu ala glu arg pro ala asn glu thr arg phe asp pro gln  
l u

The T3 protein sequence is:

FACKTANGTAIPIGGGSANVYVNLAPVNVGQNLWDLSTQIFCHNDYPETITDYVTLQRGSA  
SYPFPTTSETPRVYNSRTDKPWPVALYLTPVSSAGGVAIKAGSLIAVLILRQTNNYNSDDFQ  
CDVSA

The T7 protein sequence is:

FACKTANGTAIPIGGGSANVYVNLAPVNVGQNLWDLSTQIFCHNDYPETITDYVTLQRGSA  
SYPFPTTSETPRVYNSRTDKPWPVALYLTPVSSAGGVAIKAGSLIAVLILRQTNNYNSDDFQ  
CDVSARDVTVTLPDYRGSVPIPLTVYCAKSQNLGYYLSGTHADAGNSIFTNTASFSPAQGVG  
GAVGTSAVSLGLTANYARTGGQVTAGNVQSIIGVTFVYQ

2000-04-14 10:00:00